DOCUMENT RESUME

ED 152 829 TM 007 025

AUTHOR	Clarke, S. C. T.; And Others
TITLE	General Report on Edmonton Grade III Achievement. 1956-1977 Comparisons.
INSTITUTION	Alberta Advisory Committee on Educational Research, Edmonton.
SPONS AGENCY	Alberta Dept. of Education, Edmonton.
PUB DATE	Dec 77
NOTE	50p.; Page 8 is marginally legible
EDRS PRICE	MF-$0.83 HC-$2.06 Plus Postage.
DESCRIPTORS	*Academic Achievement; *Achievement Gains; Achievement Tests; Arithmetic; *Basic Skills; *Comparative Testing; Foreign Countries; *Grade 3; Item Analysis; Language Skills; Literature Reviews; Performance Factors; Primary Education; Reading Achievement; Scores; *Testing Programs; Test Items; Test Validity
IDENTIFIERS	*Alberta (Edmonton); California Achievement Tests; California Short Form Test of Mental Maturity; Edmonton Public Schools System; Gates Advanced Primary Reading Test

ABSTRACT

To compare achievement standards of 1977 with those of 1956, three tests were administered in their original form to all third graders in a large school district. Approximately 3500 students in 1956 and 4500 in 1977 were administered the California Achievement Tests, the California Short Form Tests of Mental Maturity, and the Gates Advanced Primary Reading Tests--Type I Word Recognition and Type II Paragraph Reading. Because many changes had occurred in the intervening years, supervisors, test experts, and all third grade teachers were invited to indicate test items which they deemed inappropriate. Seventeen items involving 18 scores out of a possible 285 were so identified: 12 out of 80 in arithmetic, 6 out of 165 in reading and none in language. All of the test items, including those items identified as inappropriate, were administered. A comparison of scores, 1956 vs 1977, was tabulated for all items, and for the items remaining after elimination. Results indicated that pupils did not find as many items inappropriate as expected; this, coupled with the fact that relatively few items were identified as inappropriate in the first place supports the assumption that the test items were valid for both groups. The 1977 scores were, on the average, higher than the 1956, the only exception being the 1% decline in the California Arithmetic. The greatest gain occurred in the Gates Word Recognition (6%), followed by California Language (5%), California Test of Mental Maturity (5%), California Reading (3%) and the Gates Paragraph Reading (2%). Spelling accounted for the increase in language scores. Taking the many factors influencing school achievement into account, the 1977 third graders performed better than their predecessors. The authors recommend the recurring publication of this type of data. (CP)

ED152829

U S DEPARTMENT OF HEALTH, EDUCATION & WELFARE
NATIONAL INSTITUTE OF EDUCATION

THIS DOCUMENT HAS BEEN REPRODUCED EXACTLY AS RECEIVED FROM THE PERSON OR ORGANIZATION ORIGINATING IT. POINTS OF VIEW OR OPINIONS STATED DO NOT NECESSARILY REPRESENT OFFICIAL NATIONAL INSTITUTE OF EDUCATION POSITION OR POLICY

"PERMISSION TO REPRODUCE THIS MATERIAL HAS BEEN GRANTED BY

W. Eddy

TO THE EDUCATIONAL RESOURCES INFORMATION CENTER (ERIC) AND USERS OF THE ERIC SYSTEM."

TM

TM007 025

Planning and Research

Edmonton Grade III Achievement Study

1956 - 1977 Comparisons

Alberta
EDUCATION
December, 1977

General Report on

Edmonton Grade III Achievement

1956 - 1977 Comparisons

A Study Conducted for

The Alberta Advisory Committee on Educational Studies

by

S.C.T. Clarke, V. Nyberg, W. H. Worth

The University of Alberta

under Contract to Alberta Education, Edmonton, Alberta, Canada

Approved by

The Edmonton Public School Board

ADVISORY COMMITTEE

H. Mosychuk (Edmonton Public School Board)
H. Leslie (Edmonton Local, Alberta Teachers' Association)
C. Rhodes (Department of Education)
H. Sherk (Department of Education)
M. Young (Edmonton Local, Alberta Teachers' Association, Chairperson)

EDMONTON GRADE III ACHIEVEMENT: 1956-1977 COMPARISONS

FOREWORD

This study was commissioned by the Minister of Education, the Honorable Julian Koziak, in the spring of 1977. It was funded through the Planning and Research Branch, Department of Education. The study was conducted for the Minister and for the Alberta Advisory Committee for Educational Studies by S.C.T. Clarke, V. Nyberg, and W.H. Worth of The University of Alberta. This study is linked to a companion study on ALBERTA GRADE III ACHIEVEMENT which deals with pupils across the whole province.

Public concern about standards of school achievement was the background for the decisions which culminated in the study. However, the foresight of G.M. Dunlop, the Chairman of the Department of Educational Psychology of the Faculty of Education, The University of Alberta, in the 1950's made the study possible. He sparked the development of the Alberta Advisory Committee for Educational Research, which in 1956 sponsored a testing program. That program provided the benchmarks in the Edmonton Public Schools which made the present study possible.

There were two major strengths of the study, in its comparison of school achievement over a span of twenty-one years. One was that a total population of students was involved. The second was that because the original test booklets, with student responses, were preserved, item by item comparisons were possible. Thus, only items deemed appropriate were included in the comparisons.

The study provided soundly-based answers to the question: how well are Grade III students achieving in the 3 R's in 1977 in comparison with 1956 as measured by the tests used. It provides answers which the public wanted. The understandable and solid data of the study should be reassuring.

The time and effort of the students, teachers, and others involved in the study will have been well invested if the public's concerns about school achievement at this level are allayed.

A Technical Report on the Edmonton Grade III Achievement: 1956-1977 study which includes the Overview and Summary from this report, and many tables and figures, has been produced in limited numbers. It is prepared for those who might replicate the study, curriculum designers, and the like. It is available in the Department of Education and from the library of The University of Alberta.

S.C.T. Clarke
V. Nyberg
W.H. Worth

EDMONTON GRADE III ACHIEVEMENT: 1956-1977 COMPARISONS

ACKNOWLEDGMENTS

Many individuals, groups, and organizations contributed to the success of this study. First in time sequence and in importance of contribution was the late G. M. Dunlop, who, in the 1950's pressed for the development of an organization to support and to conduct educational research. In addition, he sparked Study #8 of the Alberta Advisory Committee on Educational Research, the 1956 Edmonton Study of Achievement. The part played by the individuals named as associated with that study (G.M. Dunlop, S.C.Clarke, R.S. MacArthur, R.C. Harper, and W. H. Worth) is also acknowledged, as is the part played by the Edmonton Public School Board of that day, and the officials and members of the Edmonton Elementary School Local of the Alberta Teachers' Association. Special recognition must be accorded to the students who wrote the tests and to the teachers and others who administered them in 1956.

Eighteen cartons of the original test booklets, as written on by the Grade III students of the day, were preserved in the Faculty of Education. The contribution of Deans H.T. Coutts and M. Horowitz, and Acting Dean F. Enns and later Dean W. H. Worth in continuing to store these bulky tests, is acknowledged.

When the possibility of a replication was considered, the first step was to ascertain whether the tests which were originally used were available. All save one were out of print. Appreciation is expressed to Teachers' College Press and McGraw-Hill/CTB for permission to reprint the required tests.

A number of persons on the administrative staff of the Edmonton Public Schools supported the idea of replication, despite concerns about the amount of student time required and the validity of old tests. Among these were H. Mosychuk, D. Vinge, T. Blowers, N. Marchak, and S. Michaels. Officials of the Edmonton Local of the Alberta Teachers' Association supported the proposal: W. Sommerfeld, N. Somerville, J. MacLaren, M. Young, and P. Cox. As the proposal was considered by various individuals and groups, it was put forward by the administration with a positive recommendation. The part played by the superintendent and secretary-treasurer, M. Strembitsky, in these processes is acknowledged. At the November 15,1976 meeting of the Edmonton Public School Board the matter was debated fully and approval was unanimous. Grateful acknowledgement

of this favorable decision is made to J. Falconer (Chairman) and trustees M. Binder, Mrs. S. Forbes, Mrs. C. Ford, H. Jamieson, V. Johnson, and E. Lund.

From the outset encouragement was given by officials of the Department of Education, particularly W. Duke, J. Reid, and H. Sherk. The research grant through the Planning and Research Branch, Department of Education and its receipt by the Alberta Advisory Committee on Educational Studies is gratefully acknowledged. A special note of appreciation is expressed to the members of the advisory committee named on the front piece for their valuable suggestions and advice.

The many students, teachers, and principals who were involved in the testing program made the study possible. The time and effort they devoted to the task is gratefully acknowledged.

Special mention must be made of the part played by D. Blackmore who performed most of the tasks in the preparation and analysis of the data. His work was at the heart of the study, and is acknowledged with appreciation. The contribution of the Faculty of Education in providing working space and storage space for tests for twenty years, is acknowledged. Dean W.H. Worth's ready approval is deeply appreciated.

N. M. Purvis, Professor, Department of Elementary Education, The University of Alberta, was associated with the study as one of the research team until his untimely death March 29, 1977. His initial support, encouragement, and help is gratefully acknowledged.

Special mention must be made of the efforts made by J. Totten, Manager, Guidance Centre, Faculty of Education, University of Toronto, to provide the 1977 revision of the Gates MacGinitie reading test. In the end, difficulties about copyright made this impossible. The extra effort made by Mr. Totton is deeply appreciated. Similarly, G. Clark and S. Moore of McGraw Hill, Toronto office, were most helpful in supplying the latest versions of the California Achievement and California Mental Maturity tests which were used. The part that these persons played in making the study possible is acknowledged with gratitude.

A number of workers assisted the study in testing, coding, packaging, and the numerous other tasks associated with the study. Included were M. Mazeppa, M. Kugler, and members of the staff of the Special Sessions office of The University of Alberta: L. Carswell, M. Nesbitt, S. Boychuk, D. Cathcart, and P. Quan. Others employed specially to help with the study included C. Alton, J. Bishop, A. Blanes, J. Clifton, R. Coote, B. Dick, S. Friesen, R. Gerard, V. Lockwood, L. Mackeigan, G. Masson, C. McFarland, D. Neitsch, B. Panas, P. Stewart, G. Stranks and N. Wood.

The Department of Education, Alberta Education, punched all of the data onto computer tapes. This proved to be a sizeable contribution to the study which is gratefully acknowledged.

LIST OF TABLES

	Page
Difficulty and Discrimination of Items Deemed Inappropriate	28
Comparisons of Scores, 1956-1977, Items Remaining after Eliminating Items Deemed Inappropriate	30
Comparison of Scores, 1956-1977, All Items Included	31
Age of Students by Year and Sex	32

EDMONTON GRADE III ACHIEVEMENT STUDY

Table of Contents

	Page
FOREWORD	i
ACKNOWLEDGMENTS	iii
LIST OF TABLES	vi
CHAPTER I OVERVIEW AND SUMMARY	
Procedures	1
Findings	2
Discussion	4
CHAPTER II BACKGROUND	
Origin of the Study	6
Purpose of the Study	6
Limitations of the Study	7
CHAPTER III REVIEW OF THE LITERATURE	
Review of the Literature	8
CHAPTER IV THE 1956 EDMONTON GRADE III SURVEY OF ACHIEVEMENT	
The 1956 Edmonton Grade III Survey of Achievement	15
Study #8 Edmonton Study of Achievement	16
Minutes of the Edmonton Public School Board	17
Alberta Journal of Educational Research	18
CHAPTER V OVERVIEW OF THE 1977 REPLICATION	
Overview of the 1977 Replication	20
Publicity	22
Tests and Testing Procedures	24

Page

CHAPTER VI CURRICULAR VALIDITY

Curricular Validity 26

CHAPTER VII COMPARISONS

Comparisons 29

CHAPTER VIII DISCUSSION OF RESULTS

Discussion of Results 33

References 36

CHAPTER I

OVERVIEW AND SUMMARY

Procedures

The purpose of the Edmonton Grade III Achievement: 1956 - 1977 study was to compare achievement in reading, arithmetic, and language involving all of the students in a large school system. The tests, which were first given in 1956, were designed to test basic skills. The Gates Word Recognition test consists of 48 items, each of which shows a picture, followed by four words, one of which is the name of the object pictured. The Gates Paragraph Reading test has 32 tasks in 24 items, each involving one or more sentences which tell the student to draw a line from one part of a picture to another, or to place an X on something described in the paragraph. In order to do these tasks the student must be able to read the paragraph. The California Achievement Test (Reading) has 85 items in 6 subtests, which are designed to measure achievement in word form, word recognition, meaning of opposites, following directions, "directly stated facts", and interpretations. The California Achievement Test (Arithmetic) has 80 items in 9 subtests, designed to measure achievement in number and sequence, money, number and time, signs and symbols, arithmetic reasoning problems, addition combinations, subtraction combinations, multiplication problems, and arithmetic fundamentals problems. The California Achievement (Language) test has 40 items in three subtests, designed to measure achievement in capitalization, punctuation, and spelling (25 of the 40 items). The California Mental Maturity test is designed to measure general scholastic ability.

These tests, which were administered to all of the Grade III students in the Edmonton Public School system about the end of May, 1956, were all reprinted. Using the same tests, about the end of May, 1977, all of the Grade III students were again tested, using procedures which were, insofar as was possible, identical with those used in 1956.

Because many changes have occurred in the intervening years, teachers were invited to indicate any test items which they deemed inappropriate. In addition, supervisors in elementary education from the school system and experts from The University of Alberta were also invited to identify such items. Although students attempted to do all of the items, when results were computed those identified as unsuitable by the teachers, supervisors, and experts were eliminated for purposes of the comparisons.

A total of 17 items deemed inappropriate were identified: four from the Gates Word Recognition test, one (with two parts) from the Gates Paragraph Reading test, and twelve from the

California Achievement Test (arithmetic). The identification of a greater proportion of arithmetic items is not surprising because it is generally recognized that achievement in many aspects of arithmetic is closely related to what is included in the curriculum or actually taught.

The original tests used by the students in 1956 had been retained. These were rescored so that the results could be placed on computer tapes, along with the 1977 results. This procedure also ensured uniformity in scoring and permitted comparisons, item by item, for the achievement questions in the five achievement tests. It also permitted age comparisons.

A summary of the main features of the data collected follows:

1. Identical tests were used in 1956 and in 1977.
2. All of the Grade III students in a large school district were tested. The study did not rely on samples.
3. The numbers of students were large: approximately 3500 in 1956 and 4500 in 1977.
4. The tests used were standardized tests of basic skills in reading, arithmetic, and language.
5. The procedures used in administering the tests were as close to being identical as they could be made.
6. Teachers, supervisors, and experts were invited to identify test items which they deemed inappropriate. Seventeen items involving 18 scores out of a possible 285 were so identified, 12 out of 80 in arithmetic, 6 out of 165 in reading, and none in language.
7. Students did all of the tests, including the items identified as inappropriate.
8. The 1956 tests were rescored, using the same persons and procedures as were used for scoring the 1977 tests.
9. Ages of the students were available for both years.
10. Item by item results for each student for each test were recorded on computer tape.

Findings

The average score on the Cates Word Recognition test, after eliminating 4 items deemed inappropriate, made by 3540 Grade III students in 1956 was 35.35. The corresponding score made by 4430 Grade III students in 1977 was 37.41, an increase of 2.06 scores

or 6%. The average score on the Gates Paragraph Reading test, after eliminating one item (with two scores) which was deemed inappropriate, made by 3569 Edmonton Grade III students in 1956 was 22.27. The corresponding score made by 4430 students in 1977 was 22.66, an increase of 0.39 or 2%. The average score on the California Achievement Test (Reading) made by 3576 Edmonton Grade III students in 1956 was 75.26. The corresponding score made by 4417 Grade III students in 1977 was 77.58, an increase of 2.32 or 3%. No items were eliminated from the California Achievement Test (reading). The average score on the California Achievement Test (arithmetic), after eliminating 12 items deemed inappropriate, made by 3577 Edmonton Grade III students in 1956 was 55.98. The corresponding score made by 4367 Grade III students in 1977 was 55.56, a decrease of 0.42 or 1%. The average score on the California Achievement Test (language) made by 3564 Edmonton Grade III students in 1956 was 24.24. The corresponding score made by 4391 Grade III students in 1977 was 25.41, an increase of 1.17 or 5%. (Spelling accounted for the increase).

The average score on the California Mental Maturity test (a test of general scholastic ability) made by 3443 Edmonton Grade III students in 1956 was 76.09. The corresponding score made by 4378 Grade III students in 1977 was 80.12, an increase of 4.03 or 5%.

The Edmonton Grade III students in 1977 were younger than their counterparts in 1956. Boys averaged 1.96 months younger, girls averaged 1.57 months younger, and in total, the 1977 Edmonton Grade III students averaged 1.72 months younger.

A summary of the main findings of the study follows:

1. In three different tests of reading, with a total score of 159 after eliminating 6 items deemed inappropriate, the Edmonton Grade III students in 1977 performed better than did their counterparts in 1956.

2. In arithmetic, the Edmonton Grade III students in 1977 did not achieve as well as did their counterparts in 1956.

3. In language, the Edmonton Grade III students in 1977 performed better than did their counterparts in 1956 with spelling accounting for the increase in scores achieved.

4. In general scholastic ability, the Edmonton Grade III students in 1977 scored higher than their counterparts in 1956.

5. The Edmonton Grade III students when tested in 1977 were younger than were their counterparts in 1956.

Discussion

There are many factors involved in a comparative study of achievement spanning 21 years. A partial list would include: television, travel, cultural opportunities, age of school entrance, average age of students, kindergarten and nursery school provisions, school promotion policies, the school time devoted to various subjects, class size, nature of class: graded, split grades, etc., teacher qualifications, system policies, textbooks (or references), library facilities, parental attitudes, and the like. The reader can speculate on how these and other factors might affect achievement as measured.

It is most important to bear in mind that reading, arithmetic, and language were measured, in this study, by the items in the tests used. If these items did not fairly and properly measure these subjects either in 1956 or in 1977 or both, the results were, to that extent, biased or impaired. However, the small number of items identified as inappropriate by teachers, supervisors, and experts provides considerable assurance that the test items for most of the tests were considered to be reasonably fair.

Those who have asserted that the students of 1977 were not as well grounded in the basics as were their predecessors, are not supported in this contention by the evidence of this study for reading and language at the Grade III level. Their assertions may be more valid with respect to comparative achievement in arithmetic.

Both teachers and experts identified a proportionately large number of items (12 out of 80) deemed inappropriate to measure achievement in arithmetic in 1977. A small change in curriculum can cause students to perform poorly on certain items. For example, if Roman numerals are not taught in the schools, a test item asking what number corresponds to Roman XIX will be failed by most, and will so reduce the test scores. Arithmetic is "curriculum bound" in this way. In one school the principal decided not to have the students do the arithmetic part of the study because the test was not appropriate to the curriculum. In another school, one teacher took the same position. Evidently there was concern about this matter. When comparisons are made over long spans of years in arithmetic, particular care must be taken to compare on only those aspects in which the curricula are identical.

The overall results of the Edmonton Grade III Achievement: 1956-1977 Comparisons study show that the schools today are doing a commendable job in relation to their predecessors in developing

student achievement. The recurrent need for this kind of public information is great. The present study provides understandable and specific answers to basic questions about school achievement. So that such information can be available in the future, enough tests were reprinted to permit another replication about the turn of the century.

A summary of the discussion of the findings follows:

1. Many factors influence school achievement. Some of these factors undoubtedly change over the years. The study was not designed to assess these factors, so the reader must speculate about their relative effects.

2. Achievement in reading, arithmetic, and language were defined for this study by the scores students made on the test items used.

3. The study provides solid evidence that Edmonton Grade III students in 1977, despite being younger than their counterparts in 1956, out-performed them in reading.

4. The study indicates that the Edmonton Grade III students in 1977, even though they were younger, were better in language, particularly spelling, than were their counterparts in 1956.

5. The study indicates that the Edmonton Grade III students in 1977 did not do as well on the items used to measure achievement in arithmetic as did their predecessors in 1956. Arithmetic is more "curriculum bound" than are reading or language.

6. Despite being younger, Edmonton Grade III students in 1977 scored higher on a test of general scholastic ability than did their predecessors in 1956.

7. Taking all factors into account, the study shows that the Edmonton Public School system is doing a good job at the Grade III level in comparison with 1956 in developing student achievement in basic skills.

CHAPTER II

BACKGROUND

Origin of the Study

Public concern about school achievement, and public pronouncements that school achievement was in a sorry state, were background factors which set the stage for the study. These pronouncements were country-wide, continent-wide, or even more broadly made. A typical example was found in the magazine Quest (May-June 1975 issue) in an article by Gratton Gray. The title of the article was revealing: Are We Raising a Generation of Dunces , and in case the reader missed the point, the sub-title spelled it out How kids are 'doing their own thing' instead of learning to read and write. And what we should be doing about it. Two paragraphs are quoted to further illustrate Gratton Gray's contentions.

> In a nutshell, we've allowed the education system to fall victim to a set of mushy, "progressive" liberal ideas which place spurious notions of equality and "liberation" above the centuries-old values of real learning.
> A truly just society will be a mobile one, and fluid. A most important method of promoting mobility is to restore the concept of excellence to our schools, instill in students the idea that "rights" must be paid for by the assumption of obligations on their part.

While further documentation of this kind of public pronouncement could readily be provided, only one example is included, to indicate the public concern. A different kind of evidence is shown in government action. Late in 1976 the Minister of Education of the Province of Alberta established a Minister's Advisory Committee on Student Achievement. These two illustrations must suffice to indicate that there was public concern.

Purpose of the Study

The purpose of the study was to compare school achievement at the Grade III level in reading, arithmetic, and language with the corresponding achievement 21 years earlier. The 3 R's come of age! The comparison was to be one which would be easily understood by the public, since a major purpose was to inform the public about

school achievement as measured by the tests used. A total population, rather than a sample, was tested. An extensive battery of tests was involved. The same tests were to be used in 1977 as were previously used in 1956. There were enough items so that ones deemed inappropriate could be eliminated from the comparisons. With these features, the comparisons should be soundly based and easily understood.

There were 165 items which were designed to test reading, 211 for arithmetic, and 41 for language (including spelling) in five achievement tests. In addition, there was two ability tests, one of which avoided reliance on words. The ability tests could provide some explanation of the findings.

Limitations of the Study

A major limitation of the study was that the test items used provided an operational definition of achievement. When this term is used, it should be thought of as "achievement as measured by the items used." The whole area of curricular validity looms large in any replication study. If the items used were more representative of the curriculum in 1956 than in 1977, to this extent the comparison was unfair. If, however, the items used were fairly and properly representative of some absolute standard of achievement in the 3 R's, regardless of then or now, the comparisons would be fair. However, the fact remains that the items used operationally define achievement.

The most important limitation was that the study was not designed to account for the findings in comparisons of achievement. It was agreed at the outset that an examination of factors which were presumed to affect achievement was beyond the scope of the present study and if desired, should be the subject of a companion study. Such a companion study did not materialize.

Another limitation was frequently expressed rather impatiently by critics who said that they were not interested in knowing how well students achieve today in comparison with 21 years ago. They were interested in whether the students of today were achieving as well as they should. Such impatience is understandable, because the originators of such statements had not thought out how it is determined how well students should be achieving.

An obvious limitation of the study is that it was restricted to all of the students of a single school system. While many might generalize the results found, strictly speaking they would apply only insofar as related variables were the same. Many changes in the school system had occurred in the 1956 to 1977 period. One of note was the age of Grade 3 students, which was lower in 1977 than in 1956, on the average by 1.72 months. Promotion policies in the system were undoubtedly different. Curricular emphases and instructional goals were likely different. One could chronicle still further various differences in the school system between 1956 and 1977.

CHAPTER III

REVIEW OF THE LITERATURE

There are two distinct aspects to the problem of comparing student achievement at widely separated points in time. One has to do with opinions reported in newspapers, periodicals, and sometimes in books, while the other aspect involves formal research studies on standards of achievement, the results of which are reported in research journals.

Typical statements of opinions maintain that the literary skills of present-day pupils are unacceptably low, and are below the skills possessed by pupils of previous years. For example, a report by a committee of the Canadian Chamber of Commerce (1975) has a heading entitled "The Nature of Deterioration" referring to the basic skills of freshmen entering Canadian universities and colleges. Under that heading it is stated that "there has been a deterioration of the basic communication and computational skills. The opinion that this has come about is almost unanimous ...." These statements are based on responses to an opinion survey received from 125 universities and community colleges across Canada.

Saint John's Edmonton Report (1976) noted that 55% of the freshmen at the University of Calgary failed a writing test administered in September, and stated further that "much evidence has emerged to suggest that Alberta's schools, along with the rest of those in North America, are producing a generation of illiterates." The Edmonton Journal (1975) reporting on developments at The University of Alberta, noted that General Faculties Council approved in principle a recommendation that applicants for admission "demonstrate acceptable competence in English on a suitable entrance examination administered by the university."

Opinions indicating dissatisfaction with pupil achievement in the area of literacy are not limited to Alberta. Gallup (1977), reporting results of the ninth Gallup poll of attitudes towards public schools in the United States, noted that, regarding a back-to-the-basics movement (the "basics" being reading, writing and arithmetic), "all groups in the population express overwhelming approval of the movement."

A Time magazine article (1977) entitled "U.S. High Schools in Trouble" reported on high school education on the basis of reporters' visits to three schools not beset with the special problems of the large urban areas. Despite superlative buildings and equipment, and multiplicity of course offerings, standards of achievement seem to be going down. The suggested cause is academic apathy as evidenced by a high absentee rate on the part of pupils, and most of all by an unwillingness to put forth

effort. Referring to a school of 1,866 students in southern Oregon the reporters state:

> As easy courses proliferate, classes in the harder subjects wither away. Calculus and Russian, two post-Sputnik specialties, are extinct. Only 35 students are braving physics this year. Few kids prefer the no-nonsense, four-year, algebra-geometry-algebra-trigonometry sequence to the simpler math courses. The attrition rate in foreign languages is so great that after the first year, students in higher courses are combined in one class.

A publication of the Alumni Association of the University of California at Los Angeles, the UCLA Monthly (1977), devoted considerable space to an article entitled "A Plague of Illiteracy". The article reported on recent developments regarding "Subject A", a non-credit remedial course in English instituted in 1898 at the University of California. An incoming student at UCLA is directed to take this subject if an unacceptable score is achieved on the English Composition Test of the College Entrance Board. In 1968, of the entering class, 63 percent were deemed deficient, but in 1975 the figure had increased to 75 percent.

A Canadian Education Association report (Stewart, 1976), announced the formation of a committee to review examinations practices across Canada. The announcement begins by stating: "There has been considerable public criticism of alleged lowering of educational standards but there has been very little information published on what is being done to evaluate students since the elimination in most provinces of province-wide examinations."

An article in Newsweek (1976) on writing standards begins with the remark: "If your children are attending college, the chances are that when they graduate they will be unable to write ordinary, expository English with any real degree of structure and lucidity." This statement complains about low standards only, but later in the article a quote from a personnel official in a large corporation implies that standards are going down: "Errors we once found commonly in applications from high school graduates are now cropping up in forms from people with four-year college degrees."

Two older publications that were critical of education of the time both identify the mid-nineteen twenties as the time when educational standards started to decline. Rudolph Flesch (1955), advocating changes in methods of teaching reading, states: "Do you know that the teaching of reading never was a problem anywhere in the world until the United States switched to the present method?" Hilda Neatby (1952) implied that education was better in previous years when she refers to "the lost generation of Canadian education, a species last heard from nearly a quarter of a century ago ....."

Occasionally a voice is raised protesting the media assertions that standards of achievement have dropped. For example, a letter to the editor that appeared in the Edmonton Journal (1976b) concluded as follows: "I am not convinced at all that students now are not more literate than they were 25 years ago when television first hit the scene. I would put their vocabulary up against those kids any day. I'm still waiting for 'evidence' to the contrary."

There is little doubt that the prevailing mood today, as expressed in news publications, is one of dissatisfaction with pupil achievement in the basic skills. It appears, however, that many of the claims that standards are now lower than in previous years are extrapolations from statements by some authority that standards are lower than might be expected. The fact remains, however, that many people believe that average performance in the basic skills has declined in recent years.

Reports of formal studies on achievement standards are encountered less frequently than articles expressing opinions. They are sometimes based on a re-administration of a departmental examination and therefore might be of somewhat questionable value. It seems logical that students honed to a state of maximum performance will do better than students 10 or 20 years later who are neither prepared for the specific content of the test nor have an emotional investment comparable to their counterparts of previous years.

One such study (Muir, 1961) involves ninth grade algebra in Alberta. The researcher's introductory remarks are of interest: "For more than a decade public education in Canada has been the object of a critical attack of unprecedented concentration. The critics have charged that the three R's have been neglected ..." For Edmonton public school samples a drop in achievement in algebra from 42% to 36% was reported over the time from 1938 to 1959.

An investigation carried out in the Red Deer Public School District in Alberta (Alberta Education, 1976) compared achievement scores in 1972 with scores of 1976 at the ninth grade level. Grade nine departmental examinations administered in the spring of 1972 were re-administered to samples of students in three junior high schools in Red Deer. Mean scores were adjusted on the basis of intelligence scores for the 1972 and 1976 groups and comparisons made. Results indicate that there were significant declines ($P<.01$) in the means for social studies, science, language arts, and reading. In mathematics a statistically non-significant increase in the mean was recorded. The SCAT test administered along with the achievement tests in 1972 was also re-administered in 1976. There was no significant difference between the means over the four-year period. This was interpreted as an indication that the decline could not be attributed to a decrease in ability nor to a more negative response-set by pupils in 1976.

One of the most ambitious longitudinal studies to be conducted in Canada involved testing grades 5 to 8 students in the St. Catharine Public Schools (Hedges, 1977). Tests of Language Arts and Mathematics were administered in 1976, and the results were compared with data from the same tests given in the same schools in 1938 and 1954.

Several factors that might account for changes in test scores over time were also studied, and some adjustments in the scores for the tests were made so as to reduce unfairness in comparisons. Some of the more important areas of change involved organization and administration, promotion policies, teacher qualifications, curriculum, socio-economic status, age-grade levels, and supervisory assistance. The conclusions summarized in the abstract to the study are noted below:

> The major conclusions, apart from those describing the effects of the four main variables are as follows:
>
> 1. The grade 8 students do consistently less well than their earlier counterparts in arithmetic computation and reasoning.
>
> 2. The grade 5 to 7 students today, based on a test of fair comparison, outperform their earlier counterparts in fundamental operations in arithmetic.
>
> 3. The grade 8 students perform about as well, or marginally better than their counterparts in the two earlier phases in paragraph reading.
>
> 4. The performance of grade 6 students in paragraph reading skills has shown a steady but small improvement over the past forty years, as documented in the three phases of this study and an independent study conducted by Hastings in 1971.
>
> 5. The ability of students in grades 5 to 8 to initiate vocabulary from descriptions, as determined by a modern test of fair comparison, is sharply improved over their earlier counterparts, and reveals a much greater range of correct alternative responses.

In Nova Scotia longitudinal studies of school achievement are routinely conducted at regular intervals. A report involving

Grade 5 students (Nova Scotia Department of Education, 1976) observes:

> "The decline in basic skills as measured by achievement test scores reported from numerous jurisdictions for Grade 5 did not occur in Nova Scotia. On the contrary, it can be said that, for the basic skills measured by the Metropolitan, the pupils in Grade 5 met or exceeded expected performance over the two years with which this study is concerned."

This study was based on use of the 1970 edition of the Metropolitan Achievement Test.

Further information regarding Grade 3 pupils in Nova Scotia was obtained from the Standards Project file through the courtesy of the Director of Research, Mr. Arthur S. MacDonald. The table of results, below, makes it possible to compare achievement from 1955 to 1974. The results seem to indicate that achievement is up in all areas tested, with the exception of mathematics computation.

NOVA SCOTIA STANDARDS PROJECT
GRADE III
GRADE EQUIVALENT SCORES
1955 - 1974

Metropolitan Achievement Test Battery, Elementary:

Test	Year	1955	1959	1962	1965	1974
Word Knowledge		2.9	2.9	3.2	3.4	3.4
Reading		3.1	3.1	3.1	3.1	3.2
Spelling		3.5	3.6	3.8	3.0	3.7
Language		3.3	3.3	3.6	3.6	3.7
Math. Comp.		4.2	4.1	4.1	4.0	4.0
Math Concepts		3.6	3.7	3.7	3.7	3.9

Similar results were reported for Grade 6 (Nova Scotia Department of Education, 1977). Decreases in scores were reported for Mathematics Computation, Mathematics Concepts, and Mathematics Problem-solving. Scores increased for tests of Word Knowledge, Reading, Spelling, and Language.

A study of Ontario Grade 13 students (Traub, and others, 1977), employed tests in English, French, Physics and Mathematics to check on achievement standards over an interval of several years. The authors concluded that Grade 13 students achieved as well in

mathematics in 1976 as they did in 1968, but that there was a decline in physics achievement over the period 1970 to 1976. It was also reported that Grade 13 students continue to enroll in courses labelled 'traditionally academic."

A study conducted by the Planning, Development and Research Department of the Saskatoon Board of Education (1976) reported results of tests given to random samples of students in grades 6, 9 and 12 in the Saskatoon school system. Communication and computational skills were assessed by means of standardized tests. The Comprehensive Tests of Basic Skills were employed at the Grade 6 level and STEP Language Arts tests were used for measuring communication skills in Grades 9 and 12. Mathematics at the Grades 9 and 12 levels was assessed through use of the Stanford TASK Test 1-A and 11-A. In addition to the standardized tests, a written composition by each student was scored for content, form, and mechanics. The report concludes "The students in the Saskatoon public schools, as represented by the more than three hundred and fifty randomly selected students from grades 6, 9 and 12 have attained a very acceptable level of competency in both the communicative and computational skills. The criterion of acceptability utilized in the standardized test portion of this study is a normative one in which test scores of Saskatoon students have been compared to scores achieved by a standardized group of American students."

One group of noteworthy studies is involved with standardized tests administered on large scales in the United States. The results of the studies indicate that means scores tended to rise over the years, but that in the late nineteen sixties the trend was reversed. The decrease in scores is often referred to as "the test score decline". Angoff (1975), speaking of the slump in the average of the Scholastic Aptitude Test scores, claims that "there is little doubt that the decline is real."

Harnischfeger and Wiley (1976) summarize the findings of several studies by stating the following "apparent facts about achievement changes" that have occurred since the 1960's:

> Scholastic Aptitude Test. SAT scores (11th and 12th grade) have dropped for males and females in both verbal and mathematical sections.
> American College Tests. ACT scores (11th and 12th grades) have dropped in English, mathematics and social science for males and females.
> Minnesota Scholastic Aptitude Test. High school junior test scores (11th grade) have declined.
> Iowa Tests of Educational Development. ITED scores have declined in all areas for grades 9 through 12 in the State of Iowa.

They also note that no changes have occurred in the ACT Science test for grades ten and eleven, the Preliminary Scholastic Aptitude Test (for eleventh graders), the Iowa Tests of Basic Skills for grade four pupils in the State of Iowa, and for the Comprehensive Test of Basic Skills for grade five reading and language, and for grade three language. Some actual increases have been found. These are reported as follows:

> NAEP Literacy for 17-year olds and writing for 9-year olds.
> ITBS State of Iowa data for Grade 3 in all areas.
> CTBS national norm data for grades 2 and 4 in reading, language and mathematics, and grade 3 in reading and mathematics.
> Stanford-Binet. Norming data between the 1930's and 1972 show a dramatic rise.
> Metropolitan Readiness Tests. These scores have risen for pre-first-graders.

It would appear that there is no substantial body of evidence dealing with changes in standards of education in Alberta nor in Canada as a whole. Part of the reason might be that neither province-wide nor national tests of achievement are in use to any appreciable extent. In the United States there is some evidence that achievement norms for high school students are now lower than in previous years for some subject areas. In the lower grades, then, it seems likely that norms are actually higher than those of past years. This conclusion is aptly summarized by Harnischfeger and Wiley (1976) who note that declines in achievement are general in grades 5 through 12, but that "there is no evidence for decline at younger ages and in lower grades (grades 1, 2, 3, and 4), but perhaps there are increases.

The primary purpose of this study of Grade III achievement is to compare achievement standards of 1977 with those of 1956. Speculation as to causes of change are outside the frame of reference of the study, however, it should be noted that achievement comparisons at higher grade levels are inevitably more subject to the influences of curriculum changes than is the case in the lower grades. Writing skills may be a notable exception: the ability to write a clear sentence or paragraph, with words correctly spelled and punctuation properly used, is surely an important objective of any secondary school curriculum in English. This exception apart, it is clear that reading skills have high priority in the curriculum of the first three or four school grades, and therefore provide a basis for reasonable comparisons, followed by writing and arithmetic skills. The latter is probably more subject to curriculum changes than the other fields of study.

CHAPTER IV

The 1956 Edmonton Grade III Survey of Achievement

The account of the original study is based on fallible memories and fragmentary records. Three sources of written reports were available: minutes and other records of the Alberta Advisory Committee on Educational Research, minutes of the Edmonton Public School Board, and articles in the Alberta Journal of Educational Research.

The British experience with sponsoring educational research following World War II was through organizations such as the Scottish Council for Research in Education and the National Foundation for Educational Research in England and Wales. The work of these bodies was well known in the fledgling Faculty of Education at The University of Alberta. The Alberta Advisory Committee on Educational Research, which was no doubt to an extent modelled after the British organizations mentioned, included in the first pages of its record book an address by Sir Godrey Thomson delivered in November 1953 to the Sixth Annual Meeting of Council (presumably the Scottish Council mentioned above). This address contains many of the elements listed in the Foreword to the Alberta Advisory Committee on Educational Research record book. These elements which were claimed to be necessary for properly organized research called for an advisory committee, expert advisors, a project director, extended research (over five years or more), support staff, a research library, and an adequate research budget.

G.M. Dunlop was behind the movement which eventuated in the formation of the Alberta committee. Its functions were to suggest areas of needed research, to concern itself with the financing of research, and to support the publication of research findings. A Faculty of Education Research Committee was created to plan, supervise, and conduct the research program. Listed in the first (March, 1955) issue of the Alberta Journal of Educational Research were the following officers of the Alberta Advisory Committee on Educational Research: H.E. Smith (Dean of the Faculty of Education)Chairman, G.M. Dunlop (Chairman, Department of Educational Psychology) Director, and S.C.T. Clarke (Associate Professor, Department of Educational Psychology), Assistant Director.

We have here the record of a movement of the early 1950's, which culminated in a representative committee to promote educational research, and a Faculty committee to execute it. The early records of the various research projects which were supported listed financial support and little else.

Study #8 Edmonton Study of Achievement

The early studies were numbered serially, and the financial statement of the Alberta Advisory Committee on Educational Research for September 30, 1955 to April 6, 1956 shows receipts as follows: "Grant from Edmonton A.T.A. Local for Grade III Study $200.00, Grant from School Board $500.00, Grant from University Research Committee for Grade III Study $1,600.00". The same report, under disbursements lists "Edmonton Study of Achievement of Grade III #8 $1,500.00". The minutes of the April 13, 1956 meeting, as recorded by S.C.T. Clarke, acting secretary, stated that "the Edmonton Study of Achievement in Grade III is just starting. This is a large scale study which will be reported on later". The November 27, 1956 minutes included statements of receipts and disbursements from October 1, 1956 to September 30, 1957. These times overlap with the ones previously quoted. This later statement listed a receipt of $300.00 from the "Edmonton Elementary Teachers' Association", and "University Research to Grade III #8 $1,620.00". On the disbursements side, an amount of $9,120.00 was shown, with four other studies receiving $3,500.00, and the following being listed: "Edmonton Study of Achievement #8 $1,500.00, Edmonton Study of Achievement #8 $1,500.00, Edmonton Study of Achievement #8 $1,000.00, University Research to Grade III #8 $1,620.00". These sums account for the total.

It is not clear whether the first $1,500.00 granted was additional to, or one of the $1,500.00 amounts which were later recorded. Because of the total expenditure reported, the latter is assumed to be the case. If this is so, $5,620.00 was allocated to study #8, made up of $1,620.00 from the University Research fund, $500.00 from the Edmonton Public School Board, $200.00 from "The Edmonton A.T.A. Local", and the balance of $3,300.00 from other funds of the Alberta Advisory Committee on Educational Studies. It should be borne in mind that The Edmonton Elementary Teachers' Association made a further grant of $300.00. This was not listed as earmarked for Study #8.

A subsequent report lists $56.18 unexpended for Study #8, which therefore cost $5,563.82. The records of the Committee give no indication of the nature of Study #8 other than it was an Edmonton Study of Grade III Achievement.

Minutes of the Edmonton Public School Board

The record itself is so revealing that it is quoted in full.

Superintendent of Schools, April 13th, 1956, re Proposed Survey of Achievement in Grade III.

The Superintendent reported that it is proposed to conduct in cooperation with the Alberta Advisory Committee on Educational Research a proposed survey of achievement of all the Grade III pupils in the Edmonton Public Schools with respect to reading, language, spelling and arithmetic. The project is being sponsored by Messrs. G.M. Dunlop, S.C. Clarke, R.S. MacArthur, R.C. Harper and W. H. Worth. Financial support is being provided by affiliated bodies which include the University of Alberta Research Fund, Faculty Committee on Educational Research, and the Edmonton Elementary Local of the A.T.A.

The Superintendent recommended that approval be given to the proposed survey of achievement in Grade III; that the board contribute the services of the Grade III teachers in administering some of the tests and $500.00 towards the cost of the tests to be charged to the "guidance account".

MOVED BY Trustee D. H. Bowen:
"That the recommendation of the Superintendent of Schools for a proposed survey of achievement of all the Grade III pupils in the Edmonton Public Schools with respect to reading, language, spelling and arithmetic, with the Board to contribute the services of the Grade III teachers in administering some of the tests, and $500.00 towards the cost of the tests, be referred to the Management and Discipline Committee for consideration".

Unanimously carried.

The Management and Discipline Committee duly considered the matter, and the minute book records the following.

RE: Proposed Survey of Achievement in Grade III

That the recommendation of the Superintendent of Schools, "That the Board assist in the proposal to Survey the Achievement of the Grade III pupils in the Edmonton Public Schools with respect to reading, language, spelling and arithmetic:

(a) By contributing the services of the Grade III teachers in administering some of the tests;

(b) $500.00 towards the cost of the tests, to be charged to Guidance Account";

be approved.

Alberta Journal of Educational Research

A study on the effect of grouping on the variability in achievement at the Grade III level by Clarke (1958) used the data from the Edmonton Grade III survey. The study acknowledges this source as follows.

> In the last two weeks of May, 1956, all of the Grade III pupils in the Edmonton Public Schools were given a battery of tests as part of a large-scale study. The resulting data, for five large schools including 21 Grade III classrooms with 621 pupils were analyzed in the present study.

Of interest is the finding that grouping by reading reduced variability in language and arithmetic by 20%.

A second study by Worth (1959) on promotion vs non-promotion also used the Grade III data. The description sheds light on how the Grade III study was conducted.

> Selection and Description of Subjects
>
> Late in the 1955-56 school year preliminary data on all third grade pupils in the school system were obtained through the administration of the following standardized tests:
>
> 1. California Achievement Test, Complete Battery (reading, arithmetic, language), Primary Form CC.
>
> 2. Gates Advanced Primary Reading Test, Type 1-Word Recognition, Form 1.

3. Gates Advanced Primary Reading Test, Type 2- Paragraph Reading, Form 1

4. California Short-Form Test of Mental Maturity, Primary, S Form.

The achievement battery and the reading tests were administered by the classroom teachers following a training session with the writer. A trained research team was responsible for the administration of the intelligence test, and for the scoring of all of the tests. All tests were given in the same order, and at similar time intervals, in an effort to establish some degree of uniformity in the testing situation.

A later study by Worth (1961) sought to determine which students profited most from non-promotion. The finding was that older students of lower I.Q. profited most. The sixty-six subjects who were not promoted were part of the original Grade III study.

A number of unpublished studies, mostly M. Ed. theses, also used data from the Grade III study.

The test booklets were preserved for twenty years by the Faculty of Education, University of Alberta. They were labelled "Grade III, Clarke" and in 1976 fourteen cartons of tests were moved to his home. It was discovered that there were tests for 50 of the 58 schools. Later, the booklets for one other school were used up in providing illustrative material as part of approval for the study, and still later the test booklets for the missing schools were found. There were original test booklets for approximately 3,680 of the approximately 3,700 students tested. The figures are approximate because with five different test booklets, the numbers varied slightly. There was also a folder which provided frequency distributions (tallies) for all schools for all tests, the frequency distribution of total scores, means and standard deviations, and grade scores.

The data base from 1956 was from a sizeable total population of students. Five achievement tests were administered by teachers under conditions which were as standardized as was possible to arrange. Two ability tests were administered by trained testers from the university. All tests were scored by the university. The test booklets actually used by the students were retained, and were available in 1977. The frequency distribution of scores by schools, and for all students, for each test were available, as were derived statistics. This was the data base.

CHAPTER V

OVERVIEW OF THE 1977 REPLICATION

With the detailed data on hand from the 1956 study, and the temper of the times as described in Chapter II, it was agreed by many that a replication which would provide some hard data would be in order. Initially, the sponsorship of the original organizations behind the 1956 study was sought. The teachers in 1977 were organized in one local, in contrast with the three of 1956. Officials of the local were not in a position, because of finances, to provide financial support. However, they agreed to participate in and jointly sponsor a replication.

The Edmonton Public School Board was asked to participate in and jointly sponsor the replication. At a meeting held on November 16, 1976, the Board had before it the following recommendation from the administration.

> That the Edmonton Public School Board approve the request by Dr. S.C.T. Clarke to conduct a replication study designed to examine the achievement and mental maturity of Grade III students in the Edmonton Public School system.

The matter was debated by the Board. It was indicated that no direct financial contribution from the Board was requested, but that teacher and student time, about five hours, would be involved, as well as assistance from central office personnel. The matter of a 15% sample versus all students was considered. It was agreed that while such a sample could represent all students, the public was suspicious of samples but accepted results from a total population. In addition, the 1956 study was of this nature. Amendment to insert the word "all" between "of" and "grade" was considered. Some trustees were concerned that students were being over-tested. The amendment carried. The matter of ancillary studies which might throw light on or explain the comparisons in achievement and ability was considered, but was separated out from the present study. The motion as amended was carried.

The Planning and Research Branch of Alberta Education was prepared to fund the study. Major costs were reprinting tests, handscoring of tests and coding them for computer analysis, administering ability tests, analysis of data, and the like. Officials of the Branch would have been pleased had a companion study which might have thrown light on differences in achievement been forthcoming. However, this was not proposed.

The Alberta Advisory Committee on Educational Studies (formerly the Alberta Advisory Committee on Educational Research) agreed to contract with Alberta Education through the Planning and Research Branch to be responsible for the study, which was to be conducted by personnel from The University of Alberta: S.C.T. Clarke as project director, assisted by V. Nyberg and N.M. Purvis. Unfortunately, on March 29, 1977, N.M. Purvis died. This was an untimely and unexpected blow to the study team. It was with great pleasure that the remaining members were able to add W. H. Worth, Dean, Faculty of Education, The University of Alberta, to the team to replace Professor Purvis. It was particularly fitting that Dean Worth be a member of the research team, since he was one of the original 1956 researchers.

Prior to joining the study team, Dean Worth agreed that the Faculty of Education should provide storage space for records for the next twenty years, and also provided the use of a large classroom suitable for the tasks of assembling and packaging tests for distribution to schools, and for marking and coding completed tests.

Before formal arrangements were commenced, it had been ascertained, that all but one of the tests used were out of print. However, the test publishers graciously agreed to permit reprinting 15,000 copies, with provision for a modest royalty and appropriate acknowledgment. The tests which were used in the 1956 study and were reprinted were the following:

(1) California Achievement Tests Complete Battery Primary Grades 1, 2, 3 and 4. Form CC Reading, Arithmetic, Language. Devised by Ernest W. Tiegs and Willis W. Clark, Published by California Test Bureau, 5916 Hollywood Blvd., Los Angeles 28. Copyright 1950.

(2) California Short Form Test of Mental Maturity Primary Grades 1, 2, 3, 1953 S-Form. Devised by Elizabeth T. Sullivan, Willis W. Clark, and Ernest W. Tiegs. Published by California Test Bureau, 5916 Hollywood Blvd., Los Angeles 28. Copyright 1953.

(3) Gates Advanced Primary Reading Tests For Grade 2 (second half) and Grade 3. Form I. Bureau of Publications, Teachers' College, Columbia University, New York. Copyright 1942 by Arthur J. Gates.

Type I WORD RECOGNITION

(4) Gates, as above.

TYPE II PARAGRAPH READING

There was one test which was not out of print, as follows:

(5) Raven Progressive Matrices 1947. Sets A, A_B, B. Published by The Psychological Corporation, 522 Fifth Ave., New York 36, New York, under special arrangement with H.K. Lewis & Co., London, England 55-138 AS.

About 5,000 copies of each test were required for the present study. The remainder were to be kept in the event that some twenty years later someone or some group might wish to conduct a further replication. If this were done, the 1956-1977-1998 sequence might provide some very interesting results.

Publicity

At the time of the meeting of the Edmonton Public School Board publicity about the replication occurred. The November 22, 1976 Staff Bulletin of the Edmonton Public Schools which was distributed to all schools contained the following item in the Board Highlights section.

> Approved replication of 1956 study dealing with reading and writing abilities of Grade III students to be carried out during the 1976-77 school year.

The Edmonton Journal of November 17, 1976 carried an article under the headings Mental maturity, achievement to be gauged with a main heading 1956 tests for Grade 3 students. The first few paragraphs were as follows:

> Next spring all Grade 3 students in the Edmonton public school system will write a series of tests which their predecessors first wrote in 1956.
>
> The study, involving nearly 5,000 students and 250 teachers, is designed to examine and compare the achievement and mental maturity of Grade 3 students 21 years ago with those today.
>
> After considerable discussion, public school trustees voted approval of the study Tuesday night. It will be directed by Dr. S.C.T. Clarke, educational psychology professor at the University of Alberta, who proposed the replication study in May, 1975.
>
> The tests include two batteries of achievement tests and two mental ability tests, stressing proficiency in reading, language and arithmetic. They will require five hours of testing time.

An editorial in the Journal of November 19, 1956 supported the social utility of the study, commended the idea of a base of achievement, and suggested that what was needed was an extension of such studies to higher grades.

St. John's Edmonton Report of November 29, 1976 carried an interesting and colorful account, marred by some inaccuracies. Again, the first part is quoted. The headline used was Comparisons with a sub-title Local third grade to take dusty 1956 achievement test.

> With their professionalism assulted on all sides by charges that they can no longer teach children to read, write, multiply, add, subtract, or divide, four divisions of Edmonton's educational establishment have moved, with rare unanimity, to prove that their critics are wrong. Next May, some 5,000 Grade 3 pupils will take a five-hour battery of 20 year-old tests. The object: to compare their skills in the three R's with those of third graders in 1956. The testing program will be funded jointly by the Edmonton Public School Board, which recently approved the plan, the Alberta Teachers' Association, the University of Alberta and the Alberta Department of Education. Behind the project is Dr. S.C.T. Clarke, U of A educational psychologist, who says that when he discovered 14 boxes of old test booklets in his basement last spring, he realized, as he blew off the dust, that these could provide a rare opportunity for the schools to conduct a comparative study and possibly rebut their critics.

To provide general information for teachers, principals, and others who were associated with the study, a short article was prepared for the Staff Bulletin. It emphasized that the tests used in 1956 were regarded as a pool of items, from which could be eliminated items deemed inappropriate for comparisons of achievement. It spelled out what teachers were asked to do (administer tests) and what they were not asked to do (mark tests). It also called for help in locating directions used during the 1956 administration.

More detailed information was provided for principals and teachers of Grade III students. It included a statement of the purpose of the study, the history of the 1956 survey, an account of the 1977 replication and the assistance requested. This was mailed to the schools during April 1977.

The Edmonton Journal obtained copies of the April 4 Staff Bulletin with the information enclosed and the April 9, 1977, issue of the newspaper carried another story on the study. One of the researchers addressed the Division I teachers of the Edmonton Public Schools on April 13, 1977. In these and other ways, considerable publicity was provided.

Tests and Testing Procedures

A companion study Alberta Grade III Achievement using the latest available versions of the tests originally used in 1956 was planned, involving a ten percent sample of all Alberta Grade III students.

A province-wide sample had to include Grade III students from the Edmonton Public School System. To avoid overtesting Edmonton Grade III students, a matrix sampling procedure was adopted, and one of the two ability tests was dropped from each of the four Edmonton ten percent samples.

In consultation with N. Marchak, Director of Pupil Assessment E.P.S.B., the four samples were drawn by Anne Mulgren to be representative of Edmonton Grade III students. Area differences in ability level were taken into account in selecting the stratified samples. The number of students in the samples are shown on page 25.

The latest versions of the tests used were as follows:

1. California Achievement Tests, Complete Battery (reading, mathematics, language). Level 2. Form A. Copyright 1970 C.T.B. for Grades 2-4.

2. Gates-MacGinitie Reading Tests Primary C-Form 1 Vocabulary and Comprehension for Grade 3. Copyright 1964. Teachers' College Press, Columbia University.

3. California Short-Form Test of Mental Maturity, Level 1H, Grades 3 and 4. Copyright 1962 C.T.B.

4. The Raven Progressive Matrices Sets A, A_B, B, 1947.

The schedule for testing is given in Appendix 2 and is repeated here where A_o refers to the original Gates tests and A_L to the latest, B_o - C_o the original version of the California Achievement Tests and B_L - C_L the latest, D_o the original version of the California Test of Mental Maturity and D_L the latest and E_o the original version of the Raven Progressive Matrices. The tests boxed were administered by testers from the university and the achievement tests were administered by the teachers.

Groups	Test units for Edmonton Grade III Students				
Group 1 (10% sample N=461)	A_o	B_o	C_o	A_L	D_o
Group 2 (a different 10% sample N=461)	A_o	B_o	C_o	B_L	D_o
Group 3 (a different 10% sample N=466)	A_o	B_o	C_o	C_L	D_o
Group 4 (a different 10% sample N=462)	A_o	B_o	C_o	D_L	D_o
Group 5(not in samples N=2838)	A_o	B_o	C_o	E_o	D_o

The ability testers were university students. Two training sessions were held with them, which emphasized the need for clear directions and adherence to the time limits and other requirements of the test manual. Tests were scheduled for both morning and afternoon periods, but because of shorter time spans in many schools in the afternoons, tended to concentrate in the mornings. The testers took the tests with them and carried them away to be marked later.

The achievement tests were packaged by teacher and school and distributed to the schools through the Edmonton Public School System distribution system. General directions were included in the packages. The specific directions for administering the tests were photocopies of the respective parts from the appropriate test manuals.

The tests were delivered to the Edmonton Public School System May 4 and were distributed to the schools during the next week. They were returned to the school system on completion, and from there to the university. The 1956 tests were rescored, and all of the tests were coded for keypunching onto tape at the university. The test booklets were then taken to the Department of Education for keypunching. On completion of this phase, they were returned to the university for storing, so that the study might be replicated again in the future.

CHAPTER VI

CURRICULAR VALIDITY

As noted in the section on "limitations," curricular validity looms large in any replication study. Meaningful comparisons between students in 1956 and those in 1977 rest on the assumption that the test items are appropriate and fair for both groups.

To assess the soundness of this assumption each of the 224 Grade III teachers in Edmonton was asked to identify "inappropriate" items on each of the achievement tests and to give their reason(s) for doing so. Of the 206 teacher replies about the Gates Word Recognition Test, 103 identified no inappropriate items, while 154 of the 211 responding about the Gates Paragraph Reading Test discerned no inappropriate items. The same response pattern was apparent for the California Achievement Test (1950) -- 171 out of 211 did not identify any inappropriate items.

Similarly, fifteen "experts" on the teaching of language, mathematics and reading at the University of Alberta were asked for their individual opinions. The expertise of the five persons in each of the three subject-areas stemmed from their involvement in offering curriculum and instruction courses in the area to intending teachers and graduate students. Persons on the central office staff of the Edmonton Public School Board with supervisory responsibilities in these subject-areas were also requested to provide the same kind of information. Both groups also identified few inappropriate items.

The information gathered in this way was then examined to determine those test items which might be judged as being inappropriate. An item was judged as being inappropriate when it was clearly so perceived by many Edmonton teachers and when this perception was shared by experts and supervisors, as well as teachers in the provincial study.

This judgmental process gave rise to the identification of the four of forty-eight items in the Gates Word Recognition Test as being inappropriate because of obsolescence in usage or unfamiliar illustrations. Only one of twenty-four items was deemed inappropriate in the Gates Paragraph Reading Test as being contrary to the current curriculum. While twelve of 180 items on the California Achievement Test (1950) were considered inappropriate, all of them were items in arithmetic involving operations on material not customarily taught in Grade III.

These seventeen test items (one with two parts) were then subjected to the analysis summarized in Table I. This item analysis was intended to verify the judgments of inappropriateness. A distinct drop in the item difficulty figure from 1956 to 1977 may be interpreted as indicating that a certain item was not appropriate for the 1977 pupils. Similarly, a decline in the discrimination figure for an item would also imply less relevance for the 1977 group.

If a decrease of .05 is assumed to be the shift in difficulty needed for judging an item inappropriate, Table 1 shows that the initial judgments were borne out in 9 of the 17 questions as follows:

Gates Word Recognition item 24

California Arithmetic Test 3C items 8, 9, and 10

California Arithmetic Test 3E item 4 part 2

California Arithmetic Test 4I items 4, 7, and 10

The largest shifts occurred in items 9 and 10 of the California Arithmetic Test 3C which were on Roman numerals. They probably reflect a shift in teaching emphasis away from this topic.

In two of the items there was a shift of at least .05 in the difficulty in the positive direction; that is, the 1977 pupils found the items easier. These items were:

Gates Word Recognition item 44

Gates Paragraph Reading item 31 (both parts)

In only one item, Gates Word Recognition #47, was there a .05 or greater decrease in discrimination where there was no substantial drop in difficulty. This phenomenon may mirror an interest factor; that is, a few pupils of average or low achievement might have learned the meaning of the work "pugilist" through out-of-school experiences with books and/or television.

The results of this analysis suggest that the pupils did not find as many of the items inappropriate as teachers, supervisors, and experts thought they would. This finding, when coupled with the fact that relatively few inappropriate items were identified to begin with, strongly supports the assumption that the test items designed to measure communication skills were appropriate and fair for both groups. Add to this the fact that the inclusion or exclusion of the items judged as inappropriate had no influence upon the direction of the percent increase in the mean scores on any of the tests and the case for curricular validity becomes substantial. With respect to arithmetic, most of the test items were appropriate for both the 1956 and 1977 Grade III students. There were, however, a substantial number which were inappropriate for the 1977 group, while a few were also inappropriate for the 1956 group. Eliminating such items improves curricular validity but is of limited value in improving a test.

TABLE 1

DIFFICULTY AND DISCRIMINATION OF ITEMS DEEMED INAPPROPRIATE

Legend: Diff. - proportion of students who responded correctly to an item
Disc. - point biserial correlation between item response and total score on all other items in the test

Test & Item	1956		1977	
	Diff.	Disc.	Diff.	Disc.
Gates Word Recognition				
1. Item 24 - veil	0.707	0.367	0.649	0.302
2. Item 42 - arbor	0.432	0.331	0.419	0.303
3. Item 44 - dormitory	0.564	0.345	0.650	0.317
4. Item 47 - pugilist	0.232	0.104	0.259	0.041
Gates Paragraph Reading				
1. Item 31 - temperature	0.144	0.244	0.200	0.349
2. Item 31 (2nd part) "	0.214	0.363	0.288	0.327
C.A.T. Arithmetic				
1. Test 3 Sec C 8 Roman III	0.888	0.370	0.720	0.326
2. Test 3 Sec C 9 Roman VI	0.660	0.379	0.288	0.342
3. Test 3 Sec C 10 Roman XX	0.684	0.395	0.315	0.363
4. Test 3 Sec D 10 "lb." means	0.804	0.373	0.822	0.427
5. Test 3 Sec E 3 part 2 "write how"	0.816	0.355	0.830	0.315
6. Test 3 Sec E 4 part 2 "write how"	0.725	0.319	0.624	0.370
7. Test 3 Sec E 5 part 2 "write how"	0.861	0.315	0.843	0.340
8. Test 4 Sec I 4 add 4 digits	0.863	0.219	0.738	0.341
9. Test 4 Sec I 7 subt. 5 digits	0.518	0.362	0.371	0.396
10. Test 4 Sec I 8 mult. 2 x 3 digits	0.019	0.105	0.028	0.133
11. Test 4 Sec I 9 divide $4\overline{)16}$	0.554	0.290	0.562	0.352
12. Test 4 Sec I 10 divide $35\overline{)3850}$	0.090	0.196	0.018	0.126

CHAPTER VII

COMPARISONS

The main results of the study are shown in Table 2, which compares the means for each of the tests. It is readily noted that the 1977 scores were, on the average, higher than the 1956 scores. The greatest increase occurred in the Gates Word Recognition scores (6%), followed by California Language (5%), California Reading (3%) and then Gates Paragraph Reading (2%). The only decline was recorded in California Arithmetic (1%). Calculations not revealed in the tables indicate that the gain in Language was caused almost entirely by higher scores on the spelling section of the test.

Table 3 serves in a sense as a check on the inappropriateness of the items removed from the tests. It should be noted, first of all, that no items were removed from the California Reading and Language tests, therefore figures in Tables 2 and 3 are identical for these two tests. It is evident that only in the California Arithmetic test was there a change in the results because of the elimination of some items.

It would appear that the eliminating of items from the two Gates tests was not warranted, but for the California Arithmetic test the removal of items was indeed necessary.

Age figures shown in Table 4 are of considerable interest. The age of the 1977 pupils in Grade III averages 1.72 months less than that of the 1956 group. This is probably a reflection of the changes in the entrance requirements at the grade one level. Bevington (1957) states the admission requirements that applied when the 1956 Grade III class commenced school were as follows.

> All children who are six years old before the first day of September may enter on opening day. All other children whose sixth birthday comes on or before the last day of the following February may enter provided they have a mental age of 5 years 9 months.

The regulation that applied in 1974, the year the 1977 Grade III class entered school was as follows.

> A child may enter grade one at the beginning of the fall term if he has reached his fifth birthday on or before March 1 of that year. A child must enter school at the commencement of the fall term if he has attained the age of five years six months by March 1 of that year.

TABLE 2

COMPARISONS OF SCORES, 1956-1977
ITEMS REMAINING AFTER ELIMINATING ITEMS DEEMED INAPPROPRIATE

TEST	1956 Data		1977 Data		Increase in Scores (1977 scores minus 1956 scores)	Percent Increase in Mean Scores
	Number Students	Mean S.D.	Number Students	Mean S.D.		
1. Gates Advanced Primary Reading Tests WORD RECOGNITION Copyright 1942	3540	35.35 5.57	4430	37.41 4.31	2.06	6%
2. Gates Advanced Primary Reading Tests PARAGRAPH READING Copyright 1942	3569	22.27 4.14	4430	22.66 4.03	0.39	2%
TOTAL GATES READING 1 + 2		57.62		60.07	2.45	4%
3. California Achievement Tests Form CC READING Copyright 1950	3576	75.26 6.97	4417	77.58 5.93	2.32	3%
4. California Achievement Tests Form CC ARITHMETIC Copyright 1950	3577	55.98 5.97	4367	55.56 7.58	-0.42	-1%
5. California Achievement Tests Form CC LANGUAGE Copyright 1950	3564	24.24 5.35	4391	25.41 5.17	1.17	5%
TOTAL ACHIEVEMENT 3 + 4 + 5		155.48		158.55	3.07	2%
6. California Short Form Test of Mental Maturity Copyright 1953	3443	76.09 6.14	4378	80.12 6.82	4.03	5%

TABLE 3

COMPARISON OF SCORES, 1956-1977, ALL ITEMS INCLUDED

TEST	1956 Data		1977 Data		Increase in Scores (1977 scores minus 1956 scores)	Percent increase in Mean Scores
	Number Students	Mean S.D.	Number Students	Mean S.D.		
1. Gates Advanced Primary Reading Tests WORD RECOGNITION Copyright 1942	3540	37.28 6.19	4430	39.38 4.87	2.10	6%
2. Gates Advanced Primary Reading Tests PARAGRAPH READING Copyright 1942	3569	22.63 4.37	4430	23.15 4.36	0.52	2%
TOTAL GATES READING 1 + 2	———	59.91	———	62.53	2.62	4%
3. California Achievement Tests Form CC READING Copyright 1950	3576	75.26	4417	77.58	2.32	3%
4. California Achievement Tests Form CC ARITHMETIC Copyright 1950	3577	63.44 7.60	4367	61.72 9.29	-1.72	-3%
5. California Achievement Tests Form CC LANGUAGE Copyright 1950	3564	24.24 5.35	4391	25.41 5.17	1.17	5%
TOTAL ACHIEVEMENT 3 + 4 + 5	———	162.94	———	164.71	1.77	1%
6. California Short Form Test of Mental Maturity Copyright 1953	3443	76.09 6.14	4378	80.12 6.82	4.03	5%

TABLE 4

AGE OF STUDENTS BY YEAR AND SEX

	Year 1956			Year 1977			
Sex	Number	Mean Age In Months	S.D.	Number	Mean Age In Months	S.D.	1977 Younger (Months)
Male	1965	108,83	7.77	2390	106.87	5.82	1.96
Female	1878	107.32	7.05	2330	105.85	5.53	1.57
Total	3843	108.09	7.37	4720	106.37	5.59	1.72

The higher mean age for boys in 1956 remains in 1977, but the difference between the males and females is reduced from 1.51 months to 1.02 months (See Table 4). This change could well be a reflection of differences in promotion policies.

CHAPTER VIII

DISCUSSION OF RESULTS

This study presents convincing evidence that achievement in communication skills has risen at the Grade III level in the Edmonton public schools over the time period from June 1956 to June 1977. Achievement is here defined in terms of scores on the Gates Tests and the California Achievement Tests. It would be difficult to argue that the tests are somehow biased in favor of the 1977 pupils, because one has but to look at the test items to realize that they reflect school expectations in 1977 just as they did in 1956. While some of the items were eliminated because they seemed unsuitably difficult, the fact remains that they were also unsuitably difficult in 1956.

With respect to mathematics, it seems that the decrease is due to only a few items that deal with material that is now given less emphasis or not taught at all. The shifts in emphasis will no doubt be judged unsatisfactory by some sectors of society, while others will laud the changes. In this study there was no way of determining what skills were present in the 1977 group that were not present in the 1956 group, because the test was geared to the curriculum of the 1950s.

It was not the intention, in this study, to determine the causes of any changes that were found, however, some discussion seems warranted. It is almost inevitable that a reader of this report will wonder why achievement generally rose, therefore, it seems appropriate to offer a few speculations intended to stimulate reflection and research on this important topic.

Some rather profound changes have occurred in society that will undoubtedly have some effect on school achievement. The widespread (perhaps universal) use of television has influenced achievement. The nature of the effects is open to debate, but the "tube" could very well be responsible in part for the gains in communication skills, because of the great and varied verbal bombardment. If T.V. takes pupils away from homework, it could also account for the decline in mathematical skills.

There are other changes, however, that also seem important. The launching of Sputnik in the late 1950s caused a change in emphasis in education. Another fact that might be overlooked is the proximity of World War II. The 1956 Grade III pupils were born shortly after hostilities ceased. Many of the parents would have been war veterans who had their own special problems that probably would have affected the pupils in some way. The 1977 pupils' parents would be relatively free of war-related problems. The advent of permissiveness in society would probably

have some effect, perhaps only indirectly, on the pupils. Perhaps the conduct of children at age nine is not much affected by permissiveness, but the conduct of the parents might be influenced profoundly.

Finally, the recurring waves of criticism of education must not be overlooked. In 1956 a few ripples of discontent were evident, but with the launching of Sputnik in 1957 the criticism rose to tidal wave proportions. In Alberta, this culminated in the hearings of the Cameron Royal Commission.

Conditions in 1977 are somewhat similar because education again is under fire. Perhaps the recurring waves of criticism are related to social crises. In 1957 there was fear of Russian technology, while today we have more to choose from: continued inflation and unemployment; shortages, particularly of housing, natural gas, and oil; separatism; rising mental illness, crime, and drug abuse; more divorces, illegitimate births, and suicides; threats to privacy and liberty; and personal concerns such as too few friends. It may even be that criticism of education is on a cycle of its own, like sunspots or the seven-year itch! Perhaps the cycle relative to education is 20 years in duration.

In addition to changes in society at large, attention must be given to changes in schooling. Profound changes in society, such as those mentioned, are relatively easy to identify, but changes related specifically to education are not nearly so obvious. The list below provides some of the possible causes of the changes in achievement observed. The readers are invited to pick and choose their own causes of achievement changes from the following list:

1. increase or decrease in class size;
2. changes in educational philosophy;
3. shifts in subject matter emphasis;
4. decreased emphasis on testing in schools;
5. changes in pupil motivation;
6. changes in attitudes towards education by parents;
7. more (or less) availability of teaching aids and materials;
8. changes in admission policies for grade I;
9. increased availability of teaching training;
10. increased period of teacher training;
11. shifts in emphasis on homework;
12. increased use of open classrooms;
13. increased teacher specialization;
14. changes in methods of grouping pupils;
15. modifications in promotion policies;
16. greater availability of resource persons, consultants, etc.;
17. higher teacher morale;
18. lower teacher morale.

The findings of this study, that achievement at the elementary school level is as good or better than that of previous years, is consistent with conclusions given in several other reports (Nova Scotia Department of Education, 1976; Harnischfeger and Wiley, 1976; Traub and others, 1977). Educational standards at any grade level

will never be so high that improvement need not be sought; however, the fact remains that in one large school system in Alberta, standards of achievement for the most part improved over the 21-year period from 1956 to 1977.

REFERENCES

Alberta Education, Red Deer Regional Office. *Junior High School Student Outcome Study, 1972-1976.* Red Deer Public School District #104, August, 1976.

Angoff, W. Why the SAT Scores are Going Down. *English Journal,* 64, 3, 1975, pp. 10-11.

Bevington, W.G. Effects of Age at the Time of Entrance into Grade I on Subsequent Achievement. *Alberta Journal of Educational Research,* 4, No. 1, 1958, pp. 6-16.

Canadian Chamber of Commerce Education Committee. *Report on Basic Educational Skills.* Canadian Chamber of Commerce, June, 1975.

Clarke, S.C.T. The Effect of Grouping on the Variability in Achievement at the Grade III Level. *The Alberta Journal of Educational Research,* Vol. 4, #3, 1958, pp. 162-171.

Edmonton Journal. U. of A. takes Hard Line on Literacy. October 26, 1976a.

Edmonton Journal. Student Illiteracy not Proven Fact. Letter to the editor by W.P. Baerger, Stettler. October 27, 1976b.

Flesch, Rudolf. *Why Johnny Can't Read.* New York: Harper and Row, 1955.

Harnischfeger, A. and Wiley, D.E. The Marrow of Achievement Test Score Declines. *Educational Technology,* 16, 6, pp. 5-14.

FOLIO The University of Alberta Staff Bulletin. University Publications Office, The University of Alberta, Edmonton, Vol. 13, No. 16, 1976.

Gallup, George H. Ninth Annual Gallup Poll of the Public's Attitudes Towards the Public Schools. *Phi Delta Kappan,* Vol. 59, 1, September, 1977.

Hedges, H.G. *Achievement in Basic Skills.* A report funded under contract by the Minister of Education, Ontario, 1977.

Hulbert, T. A Plague of Illiteracy. *The UCLA Monthly.* UCLA Alumni Association, Vol. 7, No. 3, Jan.-Feb. 1977.

Muir, Norman D. A Comparison of the Competence in Algebra of the Grade IX Students of the Edmonton Public Schools in 1938 and 1959. *Alberta Journal of Educational Research,* 7, No. 4, 1961, pp. 175-184.

Neatby, Hilda. *So Little for the Mind.* Toronto: Clarke, Irwin and Company, 1953.

Newsweek. Why Johnny Can't Write. December, 1975, pp. 58-65.

Nova Scotia Department of Education. *Nova Scotia Standards Project, Technical Report, Grade 5, 1976.* Halifax: Nova Scotia Department of Education, 1976.

Nova Scotia Department of Education. *Nova Scotia Standards Project, Technical Report, Grade 6, 1977.* Halifax: Nova Scotia Department of Education, 1977.

Planning, Development and Research Department, Saskatoon Board of Education. *Student Achievement in Written Communication and in Computational Skills.* Saskatoon, Sask.: Saskatoon Board of Education, September, 1976.

Saint John's Edmonton Report. Alberta Universities Take Aim at Student Illiteracy, after 59% of Calgary Freshmen Flunk a New Writing Test. Edmonton, Alberta: *Saint John's Edmonton Report,* 3, No. 46, October 18, 1976, pp. 22-25.

Stewart, F. Evaluation Study Report. *Report of the Executive Director for the Year 1975-76 to the Directors of the Canadian Education Association.* Canadian Education Association, 1976.

Time. U.S. High Schools in Trouble. November 14, 1977, pp. 58-65.

Traub, R., Wolfe, R., Wolfe, C., Evans, P. and Russell, H.H. *Secondary-Postsecondary Interface Project 11: Nature of Students Vol.* 1. Toronto: Ontario Ministry of Education, 1977.

Worth, Walter H. Promotion vs. Non-Promotion: II. The Edmonton Study. *The Alberta Journal of Education Research,* Vol. V, #3, 1959, pp. 191-203.

Worth, Walter H. When is Grade Repetition Most Profitable? *The Alberta Journal of Educational Research,* Vol. VII, #4, 1961, pp. 217-222.